Anna's Blog

Updates about Anna's Archive, the largest truly open library in human history.

Copyright reform is necessary for national security

annas-archive.li/blog, 2025-01-31 — companion articles by TorrentFreak: first, second

TL;DR: Chinese LLMs (including DeepSeek) are trained on my illegal archive of books and papers — the largest in the world. The West needs to overhaul copyright law as a matter of national security.

Not too long ago, "shadow-libraries" were dying. Sci-Hub, the massive illegal archive of academic papers, had stopped taking in new works, due to lawsuits. "Z-Library", the largest illegal library of books, saw its alleged creators arrested on criminal copyright charges. They incredibly managed to escape their arrest, but their library is no less under threat.

When Z-Library faced shutdown, I had already backed up its entire library and was searching for a platform to house it. That was my motivation for starting Anna's Archive: a continuation of the mission behind those earlier initiatives. We've since grown to be the largest shadow library in the world, hosting more than 140 million copyrighted texts across numerous formats — books, academic papers, magazines, newspapers, and beyond.

Me and my team are ideologues. We believe that preserving and hosting these files is morally right. Libraries around the world are seeing funding cuts, and we can't trust humanity's heritage to corporations either.

Then came AI. Virtually all major companies building LLMs contacted us to train on our data. Most (but not all!) US-based companies reconsidered once they realized the illegal nature of our work. By contrast, Chinese firms have enthusiastically embraced our collection, apparently untroubled by its legality. This is notable given China's role as a signatory to nearly all major international copyright treaties.

We have given high-speed access to about 30 companies. Most of them are LLM companies, and some are data brokers, who will resell our collection. Most are Chinese, though we've also worked with companies from the US, Europe, Russia, South Korea, and Japan. DeepSeek admitted that an earlier version was trained on part of our collection, though they're tight-

lipped about their latest model (probably also trained on our data though).

If the West wants to stay ahead in the race of LLMs, and ultimately, AGI, it needs to reconsider its position on copyright, and soon. Whether you agree with us or not on our moral case, this is now becoming a case of economics, and even of national security. All power blocs are building artificial super-scientists, super-hackers, and super-militaries. Freedom of information is becoming a matter of survival for these countries — even a matter of national security.

Our team is from all over the world, and we don't have a particular alignment. But we'd encourage countries with strong copyright laws to use this existential threat to reform them. So what to do?

Our first recommendation is straightforward: shorten the copyright term. In the US, copyright is granted for 70 years after the author's death. This is absurd. We can bring this in line with patents, which are granted for 20 years after filing. This should be more than enough time for authors of books, papers, music, art, and other creative works, to get fully compensated for their efforts (including longer-term projects such as movie adaptations).

Then, at a minimum, policymakers should include carve-outs for the mass-preservation and dissemination of texts. If lost revenue from individual customers is the main worry, personal-level distribution could remain prohibited. In turn, those capable of managing vast repositories — companies training LLMs, along with libraries and other archives — would be covered by these exceptions.

Some countries are already doing a version of this. TorrentFreak reported that China and Japan have introduced AI exceptions to their copyright laws. It is unclear to us how this interacts with international treaties, but it certainly gives cover to their domestic companies, which explains what we've been seeing.

As for Anna's Archive — we will continue our underground work rooted in moral conviction. Yet our greatest wish is to enter the light, and amplify our impact legally. Please reform copyright.

- Anna and the team (Reddit, Telegram)

Read the companion articles by TorrentFreak: first, second